

Análise do Conteúdo Tempo-Frequência de Frames de Weyl-Heisenberg e sua Aplicação na Geração de Dicionários Redundantes Parametrizados

Lisandro Lovisoló^{1,2}, Eduardo A. B. da Silva² e Paulo S. R. Diniz²

1. Depto. de Eletrônica e Telecomunicações – UERJ, Rio de Janeiro, RJ, BRASIL

2. Lab. de Processamento de Sinais – COPPE/UFRJ, Cx. P. 68504, Rio de Janeiro, RJ, BRASIL

[lisandro,eduardo,diniz@lps.ufrj.br]

Resumo—Este trabalho investiga o conteúdo tempo-frequência de frames. Iniciamos mostrando que a soma dos conteúdos tempo-frequenciais de todos os elementos de um conjunto de funções ser positiva é uma condição suficiente para que este conjunto gere um frame em $L^2(\mathbb{R})$. A seguir deriva-se que para frames de Weyl-Heisenberg $\{E_{mb}T_{na}g(t)\}_{n,m \in \mathbb{Z}}$ gerados a partir de uma função par $g(t)$ os máximos e mínimos de seu conteúdo tempo-frequência encontram-se em (na, mb) e $(na + a/2, mb + b/2)$, respectivamente; e que para $g(t)$ ímpar teremos os máximos localizados em $(na, mb + b/2)$ e os mínimos em $(na + a/2, mb)$. Estes resultados fornecem uma forma efetiva de gerar frames mais apertados por entrelaçamento, ao custo dobrar a cardinalidade dos frames. Os frames construídos pela abordagem apresentada são avaliados para utilização como dicionários em decomposições vorazes de sinais.

Palavras-Chave—Frames, Frame Bounds, Dicionários, Decomposições Adaptativas.

Abstract—The work presented here investigates the time-frequency content of frames. We begin by showing that the sum of the time-frequency contents of all the elements of a set of functions being always positive is a sufficient condition for this set of functions to generate a frame. It is then derived that for Weyl-Heisenberg frames $\{E_{mb}T_{na}g(t)\}_{n,m \in \mathbb{Z}}$ of an even function $g(t)$ the maxima of the time-frequency content occur at (na, mb) and the minima at $(na + a/2, mb + b/2)$; whereas for an odd function $g(t)$ the maxima occur at $(na, mb + b/2)$ and the minima at $(na + a/2, mb)$. These results indicate effective ways to interlace frames obtaining “tighter” frame bounds. The frames constructed using the proposed approach are evaluated for the use as dictionaries in a greedy decomposition framework.

Keywords—Frames, Frame Bounds, Dictionaries, Adaptive Decompositions.

I. INTRODUÇÃO

Recentemente, tem se estudado uma classe de algoritmos de análise de sinais que decompõem/representam os sinais usando as estruturas (átomos ou elementos) de um dicionário \mathcal{D} . Esta classe de algoritmos difere das transformações clássicas de sinais, que utilizam a projeção do sinal sobre todos elementos de uma base para representá-lo, ao permitir que os elementos utilizados na decomposição/representação do sinal não necessariamente formem uma base, podendo ser, até, “redundantes”. Estas técnicas têm sido utilizadas tanto para filtragem [1], [2], compressão [3]–[7], análise dos fenômenos físicos presentes no sinal [1], [4], [6], [8]–[11] e análise tempo-frequência [1], [12] entre outras, aparecendo como alternativas às técnicas clássicas de análise e síntese por transformações.

Nesta classe de algoritmos, a decomposição é obtida escolhendo-se M elementos de \mathcal{D} que fornecem uma aproximação do sinal para um critério (que em geral consiste na minimização da norma do erro de aproximação ou um número de estruturas M pré-determinado):

$$x \approx \sum_{n=1}^M \alpha_n g_{i_n}, \quad g_i \in \mathcal{D}. \quad (1)$$

Na eq. (1), i_n é o índice da estrutura de número n no dicionário, utilizada para aproximar o sinal, e α_n é o seu peso. Em qualquer abordagem de decomposição, (para exemplos ver [1], [2], [11], [13], [14]), o objetivo é representar o sinal de forma compacta e significativa. O requisito inicial é que \mathcal{D} seja completo [1] podendo desta forma aproximar exatamente qualquer sinal. \mathcal{D} deverá ser, também, suficientemente grande de forma a permitir “boas” decomposições. Logo, é desejável a utilização de um dicionário sobre-completo (redundante). Neste, diferentes combinações das estruturas que o compõem podem ser usadas para representar o mesmo sinal. Um dos problemas desta classe de algoritmos de decomposição está no projeto e na definição do dicionário \mathcal{D} . Em algumas aplicações, utiliza-se treinamento para projetar os dicionários [15]. Entretanto, dicionários obtidos por treinamento, *a priori*, não são parametrizáveis. Dicionários parametrizáveis caracterizam-se pela definição de cada g_i a partir de um conjunto de parâmetros $i = (\gamma_0, \gamma_1, \dots, \gamma_{L-1})$ onde L é a quantidade de parâmetros que definem g_i . Talvez, o mais conhecido entre os dicionários parametrizáveis seja o dicionário de Gabor [1], [12], [16]. O dicionário de Gabor utiliza estruturas Gaussianas com diferentes escalas, centros de simetria e modulações. Assim sendo, estes podem ser parametrizados pela escala, centro e frequência de modulação da estrutura. Dicionários como o de Gabor podem ser construídos a partir de Frames de Weyl-Heisenberg e de Frames de Wavelets, que veremos posteriormente. Frames são, simplifadamente, uma base sobrecompleta. A utilização de Frames para construção de dicionários é útil em aplicações de modelagem de sinais [4]–[6], e suas origens remontam a [12], [16].

Às vezes, é desejável a utilização de dicionários que permitam representar efetivamente uma dada classe de sinais. Neste caso, dicionários parametrizáveis são uma boa escolha por permitirem encontrar estruturas com uma mesma forma mas com características que dependem dos parâmetros. Para melhorar a identificação de estruturas com uma dada forma nos sinais sob análise pode-se aumentar a densidade dos parâmetros que definem as estruturas, aumentando assim a cardinalidade de \mathcal{D} , $\mathcal{K}(\mathcal{D})$. Porém, a complexidade computacional dos algoritmos que obtêm uma representação como a da eq. (1), normalmente cresce com o aumento de $\mathcal{K}(\mathcal{D})$. É então necessário encontrar meios de aumentar a cardinalidade do dicionário que permitam analisar/decompor melhor os sinais com aumento controlado do custo computacional. Neste artigo, mostramos como aumentar a cardinalidade de dicionários gerados a partir de frames de Weyl-Heisenberg de forma efetiva, em termos da taxa necessária para codificar a representação. Este aspecto está relacionado a quão “denso” é o frame no espaço de parâmetros. Porém, a noção de “densidade” nestes espaços não é bem definida. Neste trabalho, utilizamos uma medida

do conteúdo tempo-frequência do *frame*, via distribuição de Wigner-Ville, para avaliar esta densidade. Mostra-se também que esta medida está relacionada a quão “apertado” é o *frame*.

Na seção II, aspectos elementares relativos a *frames* são apresentados. Na seção III, vemos como a distribuição de Wigner-Ville pode ser utilizada no estudo do conteúdo tempo-frequência de um sinal a partir de decomposições adaptativas. Isto motiva a definição do conteúdo tempo-frequência de um *frame*. Demonstramos, então, um teorema que fornece uma condição suficiente para um conjunto de funções g_k ser *frame* de $L^2(\mathbb{R})$, a partir de seu conteúdo tempo-frequência. Entretanto, é interessante saber como escolher os passos que definem a translação e a modulação de forma a gerar um *frame* de Weyl-Heisenberg (WH) mais apertado. Na seção IV, para os *frames* WH gerados a partir de funções pares e ímpares vemos como gerar *frames* mais apertados, permitindo uma melhor análise do conteúdo tempo-frequência de um sinal quando decomposto neste *frame*. Na seção V, mostra-se como e porque um *frame* assim gerado pode ser utilizado como dicionário em decomposições vorazes e o desempenho destes *frames* nesse contexto é avaliado.

II. FRAMES

Uma seqüência de elementos $\{g_k\}_{k \in \mathbb{I}}$ é um *frame* do espaço $L^2(\mathbb{R})$ se existirem constantes $0 < A < B < \infty$ tais que [17]

$$A\|x\|^2 \leq \sum_{k \in \mathbb{I}} |\langle x, g_k \rangle|^2 \leq B\|x\|^2, \quad \forall x \in L^2(\mathbb{R}). \quad (2)$$

Os números A e B são chamados de *lower* e *upper frame bounds*, respectivamente. Define-se ainda o *frame* inverso, que fornece as fórmulas de reconstrução

$$x = \sum_{k \in \mathbb{I}} \langle x, g_k \rangle \tilde{g}_k = \sum_{k \in \mathbb{I}} \langle x, \tilde{g}_k \rangle g_k, \quad (3)$$

onde o conjunto \tilde{g}_k , $k \in \mathbb{I}$ são os elementos do *frame* inverso.

O operador do *frame* (S) é definido como $S\{x\} = \sum_{k \in \mathbb{I}} \langle x, g_k \rangle g_k$. Uma conseqüência direta da definição de S é que

$$x = \sum_{k \in \mathbb{I}} \langle x, g_k \rangle S^{-1}\{g_k\}, \quad (4)$$

onde S^{-1} é o operador inverso do *frame*.

Se $A = B \rightarrow S^{-1}\{\cdot\} = S\{\cdot\}/A$ e o *frame* é dito apertado (*tight*) [17]. Neste caso, tem-se $\sum_{k \in \mathbb{I}} |\langle x, g_k \rangle|^2 = A\|x\|^2$, $\forall x \in L^2(\mathbb{R})$. Outra definição muito comum é a de *snug frame* para os quais $A \approx B$, isto é $S^{-1}\{\cdot\} \approx S\{\cdot\}/A$.

Frames apertados têm se tornado muito populares e recentemente muita atenção tem sido dada à construção deste tipo de *frames* [18], [19].

Salientamos ainda que a união de dois *frames* quaisquer de um espaço gera um novo *frame* desse espaço.

A. Frames de Espaços Vetoriais Finitos

Num espaço vetorial \mathbb{R}^N a definição de *frame* é

$$A\|f\|^2 \leq \sum_{k \in \mathcal{K}} |\langle f, g_k \rangle|^2 \leq B\|f\|^2, \quad f \in \mathbb{R}^N. \quad (5)$$

Define-se o operador de síntese $\mathbf{T} : \mathbb{C}^K \rightarrow \mathbb{R}^N$, $\mathbf{T}\{c_k\}_{k=1}^K = \sum_{k=1}^K c_k \tilde{g}_k$, onde \tilde{g}_k são os elementos do *frame* inverso. Define-se também o operador de análise $\mathbf{T}^*\{\cdot\}$ (operador adjunto de $\mathbf{T}\{\cdot\}$) que é $\mathbf{T}^* : \mathbb{R}^N \rightarrow$

\mathbb{C}^K , $\mathbf{T}^*\{f\} = \{\langle f, g_k \rangle\}_{k=1}^K$. Logo, o operador do *frame* é

$$\mathbf{S} : \mathbb{R}^N \rightarrow \mathbb{R}^N, \quad \mathbf{S}\{f\} = \mathbf{T}\{\mathbf{T}^*\{f\}\} = \sum_{k=1}^K \langle f, g_k \rangle \tilde{g}_k. \quad (6)$$

Então, $\langle \mathbf{S}\{f\}, f \rangle = \sum_{k=1}^K |\langle f, g_k \rangle|^2$, $f \in \mathbb{R}^N$, e temos $A\|f\|^2 \leq \|\mathbf{S}\{f\}\|^2 \leq B\|f\|^2$, e tem-se $\|\mathbf{S}\{f\}\|^2 = \sum_{k \in \mathcal{K}} |\langle f, g_k \rangle|^2$. Em espaços vetoriais finitos, todas estes operadores são dados por matrizes [17] e $\mathbf{S} = \mathbf{T}\mathbf{T}^*$. Sendo ρ_i os autovalores generalizados de \mathbf{S} , tem-se que $A = \min_i \rho_i$ e $B = \max_i \rho_i$ [17]. Assim sendo, para um *frame* apertado teremos $\mathbf{S}^{-1}\mathbf{S} = \mathbf{I}_N$ (\mathbf{I}_N é a matriz identidade de ordem N), se $\mathbf{T}\mathbf{T}^* = A\mathbf{I}_N$ tem-se um *frame* apertado.

B. Construção de Frames de Weyl-Heisenberg a Partir de uma Função $g(t)$

Um conjunto de funções g_k , $k \in \{1, 2, \dots, K\}$ gerará o espaço $L^2(\mathbb{R})$ se ele constituir um *frame* [16], [17]. Há várias formas de construir *frames* em $L^2(\mathbb{R})$ a partir dos operadores de translação de a (T_a), modulação por uma senóide de frequência b (E_b) e dilatação por c (D_c) de uma função $g(t)$ [17], que são definidos por:

$$\begin{aligned} T_a (a \in \mathbb{R}) : L^2(\mathbb{R}) &\rightarrow L^2(\mathbb{R}), & (T_a g)(t) &= g(t - a); \\ E_b (b \in \mathbb{R}) : L^2(\mathbb{R}) &\rightarrow L^2(\mathbb{R}), & (E_b g)(t) &= g(t)e^{2\pi jbt}; \\ D_c (c \in \mathbb{R}^*) : L^2(\mathbb{R}) &\rightarrow L^2(\mathbb{R}), & (D_c g)(t) &= \frac{1}{\sqrt{c}}g\left(\frac{t}{c}\right). \end{aligned}$$

Um *frame* de Weyl-Heisenberg (WH) ou de Gabor é um *frame* de $L^2(\mathbb{R})$ obtido através das operações $\{E_{mb}T_{na}g(t)\}_{m,n \in \mathbb{Z}}$, onde $a, b > 0$ e $g(t)$ é uma função fixa. Este é também conhecido como *Frame* de Gabor, Sistema de Gabor ou *Frame* Fourier-janelado [1], [12], [16].

Frames em espaços vetoriais são normalmente construídos através da amostragem no tempo dos elementos de *frames* em $L^2(\mathbb{R})$ [17]. Por exemplo, o dicionário de Gabor [1], [12], [16] é constituído de *frames* WH em diferentes escalas em espaços discretos finitos, ou seja, é a união de sucessivos *frames* $\{E_{mb/c}T_{nac}D_{kc}g\}_{m,n,k \in \mathbb{Z}}$.

III. CONTEÚDO TEMPO-FREQÜÊNCIA DE FRAMES

A distribuição de Wigner-Ville (WD) é provavelmente uma das ferramentas mais utilizadas e conhecidas para a análise tempo-frequência de sinais. A WD de um sinal $x(t)$ é definida como [1]

$$WD_x(t, f) = \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) \overline{x\left(t - \frac{\tau}{2}\right)} e^{-2\pi jf\tau} d\tau, \quad (7)$$

onde a barra é conjugação complexa, e é uma “medida” da densidade de energia no tempo (t) e na frequência (f) do sinal $x(t)$ [1].

Muitas aplicações utilizam a WD da decomposição do sinal para analisar seu conteúdo tempo-frequência. Esta abordagem é apresentada em [12] usando átomos tempo-frequência Gausianos (que possuem a melhor concentração tempo-frequência possível) para decompor os sinais (para exemplos ver [1]).

Vários algoritmos de decomposição de sinais são baseados no conceito de *frames* [16], [17]. Nestes casos, o sinal é decomposto nos elementos do *frame* e seu conteúdo tempo-frequência é “inferido” desta decomposição.

Definamos o conteúdo tempo-frequência da decomposição de um sinal $x(t)$ em um *frame* \mathcal{G} como

$$\begin{aligned} WD_{x(t) \text{ em } \mathcal{G}}(t, f) &= \sum_k \langle x(t), g_k(t) \rangle WD_{g_k}(t, f) = \\ &= \|x(t)\| \sum_k c_k WD_{g_k}(t, f). \end{aligned} \quad (8)$$

Consideremos que todos os elementos do *frame* possuem norma unitária, então $c_k = \left\langle \frac{x(t)}{\|x(t)\|}, g_k(t) \right\rangle \leq 1$, e para qualquer sinal $x(t)$ teremos

$$WD_{x(t) \text{ em } \mathcal{G}}(t, f) \leq \|x(t)\| \sum_k WD_{g_k}(t, f). \quad (9)$$

O conteúdo tempo-frequência de um *frame* \mathcal{G} é dado por

$$WD_{\mathcal{G}}(t, f) = \sum_k WD_{g_k}(t, f). \quad (10)$$

O Teorema III.1 mostra uma condição suficiente sobre $WD_{\mathcal{G}}(t, f)$ para que qualquer conjunto de funções $g_k(t)$ seja um *frame* de $L^2(\mathbb{R})$.

Teorema III.1 *Uma condição suficiente para que uma sequência de elementos $\mathcal{G} = \{g_k\}$ seja um frame em $L^2(\mathbb{R})$ é*

$$0 < WD_{\mathcal{G}}(t, f) < \infty, \quad \forall(t, f), \quad (11)$$

onde $WD_{\mathcal{G}}(t, f) = \sum_k WD_{g_k}(t, f)$.

Demonstração: A WD satisfaz a relação de Moyal $\langle WD_x, WD_g \rangle = \int_t \int_f WD_x(t, f) WD_g(t, f) df dt = |\langle x, g \rangle|^2$, a WD também conserva a energia $\|x\|^2 = \int_t \int_f WD_x(t, f) df dt$. Aplicando estas propriedades na definição de *frame* (eq. (2)) pode ser visto que

$$\begin{aligned} A \int_t \int_f WD_x(t, f) df dt &\leq \sum_{k \in \mathbb{I}} \langle WD_x, WD_{g_k} \rangle \\ &\leq B \int_t \int_f WD_x(t, f) df dt. \end{aligned} \quad (12)$$

Logo, $\forall(t, f)$ tem-se que $A \int_t \int_f WD_x(t, f) df dt \leq \int_t \int_f WD_x(t, f) WD_{\mathcal{G}}(t, f) df dt$ e $B \int_t \int_f WD_x(t, f) df dt \geq \int_t \int_f WD_x(t, f) WD_{\mathcal{G}}(t, f) df dt$. Então, é suficiente para que a expressão (12) seja válida que

$$A \leq WD_{\mathcal{G}}(t, f) \leq B \quad \forall(t, f). \quad (13)$$

Portanto, se $0 < WD_{\mathcal{G}}(t, f) < \infty$, $\forall(t, f)$, obtém-se um *frame*. ■

Este resultado mostra que se o somatório das WDs de todos os elementos de um conjunto g_k for limitado e positivo, então este conjunto é um *frame* de $L^2(\mathbb{R})$. Como este resultado surge diretamente das propriedades da WD, é válido também em espaços discretos.

Corolário III.1 *A eq. (12) mostra que é suficiente para que um frame seja apertado ($A = B$) que seu conteúdo tempo-frequência seja constante.*

Exemplo III.1 *Considere um frame no espaço de dimensão N par formado pelos vértices do cubo N -dimensional. Isto é, seus elementos são $\mathbf{g}_k = [\mathbf{g}_k[0], \dots, \mathbf{g}_k[N-1]] = [\pm 1, \dots, \pm 1]$. Pode-se mostrar que para estes frames teremos*

$WD_{\mathcal{G}}(n, k) = N^2$, ou seja, seu conteúdo tempo-frequência é constante, implicando um *frame* apertado. Podemos ver ainda que, para estes frames, $\mathbf{T}\mathbf{T}^* = N^2 \mathbf{I}_N$. Isto também mostra que este é um *frame* apertado e \mathbf{S} (ver eq. (8)) é uma matriz identidade multiplicada por $A = N^2$.

IV. CONTEÚDO TEMPO-FREQÜÊNCIA DE FRAMES WEYL-HEISENBERG

Para um *frame* WH \mathcal{G} tem-se

$$WD_{\mathcal{G}}(t, f) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} WD_g(t - na, f - mb). \quad (14)$$

Operações sobre a eq. (14) fornecem [20]

$$WD_{\mathcal{G}}(t, f) = \frac{1}{b} \sum_{m \in \mathbb{Z}} e^{-2\pi j f \frac{m}{b}} \sum_{n \in \mathbb{Z}} g(t + p_{n,m}) \overline{g(t + q_{n,m})}, \quad (15)$$

onde $p_{n,m} = -na + \frac{m}{2b}$ e $q_{n,m} = -na - \frac{m}{2b}$. Pode se derivar da eq. (15) que

$$\frac{\partial WD_{\mathcal{G}}(t, f)}{\partial t} = \frac{1}{b} \sum_{m \in \mathbb{Z}} e^{-2\pi j f \frac{m}{b}} \quad (16)$$

$$\begin{aligned} &\sum_{n \in \mathbb{Z}} \left[\frac{\partial g(t + p_{n,m})}{\partial t} \overline{g(t + p_{n,m})} + g(t + p_{n,m}) \frac{\partial \overline{g(t + q_{n,m})}}{\partial t} \right] \\ \frac{\partial WD_{\mathcal{G}}(t, f)}{\partial f} &= \frac{1}{b} \sum_{m \in \mathbb{Z}} -2\pi j \frac{m}{b} e^{-2\pi j f \frac{m}{b}} \sum_{n \in \mathbb{Z}} g(t + p_{n,m}) \overline{g(t + q_{n,m})} \end{aligned} \quad (17)$$

A. Máximos e Mínimos de $WD_{\mathcal{G}}(t, f)$ para Frames WH Gerados a partir de Funções Pares

Para obter $\max_{(t,f)} WD_{\mathcal{G}}(t, f)$ e $\min_{(t,f)} WD_{\mathcal{G}}(t, f)$ de um *frame* WH, basta analisar $WD_{\mathcal{G}}(t, f)$ na região $(t, f) \in [0, a) \times [0, b)$, já que $WD_{\mathcal{G}}(t, f)$ é 2D-periódica. Se $g(t)$ é simétrica, então pode-se mostrar, a partir das eqs. (16) e (17), que $\frac{\partial WD_{\mathcal{G}}(t, f)}{\partial t} = 0$ e $\frac{\partial WD_{\mathcal{G}}(t, f)}{\partial f} = 0$ para $(t, f) = (0, 0)$ ou $(t, f) = (\frac{a}{2}, \frac{b}{2})$. Neste caso podemos mostrar ainda que $WD_{\mathcal{G}}(0, 0) \geq WD_{\mathcal{G}}(\frac{a}{2}, \frac{b}{2})$ e tem-se que os máximos de $WD_{\mathcal{G}}(t, f)$ ocorrem em $(t = na, f = mb)$ e os mínimos em $(t = na + a/2, f = mb + b/2)$.

B. Máximos e Mínimos de $WD_{\mathcal{G}}(t, f)$ para Frames WH Gerados a partir de Funções Ímpares

Se $g(t)$ for anti-simétrica pode-se mostrar, de maneira similar à da subseção anterior, que os máximos de $WD_{\mathcal{G}}(t, f)$ ocorrem em $(t = na, f = mb + b/2)$ e os mínimos em $(t = na + a/2, f = mb)$.

C. Intercalando Frames WH

Os resultados acima indicam que para aumentar a capacidade de localização, aumentando a densidade no plano tempo-frequência, de um *frame* WH gerado a partir de uma função simétrica $g(t)$, deve-se empregar a união de *frames* gerados a partir de $E_{mb} T_{na} g(t)$ e $E_{mb + \frac{b}{2}} T_{na + \frac{a}{2}} g(t)$. Este resultado é interessante para a geração de *frames* a partir de átomos exponenciais bilaterais, átomos Gaussianos ou quaisquer outros átomos simétricos. Em [21] um resultado similar é apresentado utilizando o conceito de *frames* Grassmannianos, nos quais os

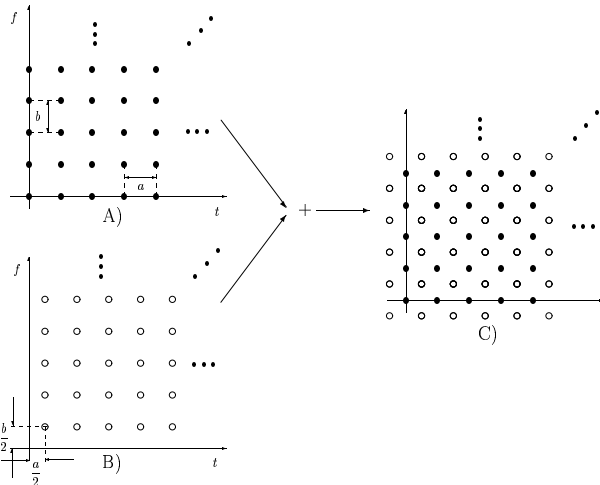


Fig. 1. Intercalando frames WH.

elementos são escolhidos de forma a minimizar seus produtos internos. Isto é o que se faz ao se situar os máximos, no plano tempo-freqüência, dos novos elementos nos pontos de mínimo de $WD_{\mathcal{G}}(t, f)$, já que pela relação de Moyal há uma relação entre o produto interno de duas funções e seu produto interno no plano tempo-freqüência. A intercalação de frames é apresentada na Fig. 1: 1.A) mostra a localização dos máximos de $WD_{\mathcal{G}}(t, f)$ no plano tempo-freqüência, 1.B) o deslocamento do frame através da translação de $a/2$ e da modulação por $b/2$ de todos os seus elementos, e 1.C) a união dos dois.

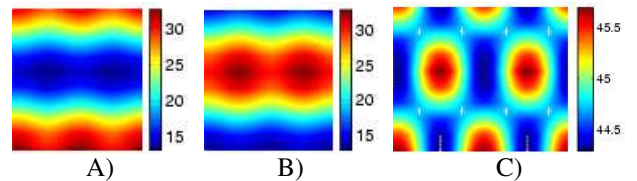
Empregando a união de frames sugerida acima obtém-se $WD_{\mathcal{G}}(na, mb) = WD_{\mathcal{G}}(na + \frac{a}{2}, mb + \frac{b}{2})$. Se podemos gerar sistemas de Gabor $E_{mb}T_{na}g(t)$ e $E_{mb+\frac{b}{2}}T_{na+\frac{a}{2}}g(t)$ tais que sua união gere um $WD_{\mathcal{G}}(t, f)$ com menores "oscilações", então obteremos um frame mais apertado. Se conseguimos fazer $WD_{\mathcal{G}}(t, f)$ constante então o frame resultante é apertado.

O mesmo raciocínio pode ser aplicado para átomos anti-simétricos (trocando as posições dos pontos de máximo e de mínimo pelas correspondentes).

Exemplo IV.1 A Fig. 2.A) apresenta o conteúdo tempo-freqüência de $WD_{\mathcal{G}}(t, f)$, de um frame $E_{mb}T_{na}g(t)$ onde $g(t)$ é uma Gaussiana $g(t) = e^{-t^2/2\sigma^2}$, com $\sigma^2 = 1/40$, $a = .25$ e $b = 2.5$, na região $[0, 2a) \times [0, 2b)$ (estes dados foram obtidos considerando uma taxa de 100 amostras por unidade de tempo). A Fig. 2.B) apresenta o mesmo, para o frame $E_{mb+b/2}T_{na+a/2}g(t)$. Por sua vez, a Fig. 2.C) apresenta o resultado da união dos dois frames. Repare que, ao realizar a união, o máximo e o mínimo de $WD_{\mathcal{G}}(t, f)$ tornam-se mais próximos. O que sugere a obtenção de um frame mais apertado.

V. DICIONÁRIOS FORMADOS POR FRAMES INTERCALADOS

Nesta seção, o desempenho de dicionários construídos através da união de frames proposta é avaliado quando estes são utilizados em algoritmos de decomposição vorazes. Os algoritmos de decomposição vorazes, escolhem, a cada passo,


 Fig. 2. Exemplo de intercalação de frames WH no intervalo $(t, f) \in [0, 2a) \times [0, 2b)$ – o eixo horizontal é o tempo e o eixo vertical é a freqüência.

o elemento do dicionário que melhor representa o sinal considerando um dado critério de erro. O átomo escolhido é então escalado e subtraído do sinal, gerando um resíduo que deve ser aproximado no próximo passo. Este processo é iterado até que um limiar de erro para a aproximação ou um determinado número de passos seja atingido.

Um destes algoritmos é o *Matching Pursuits* (MP) [12]. O MP obtém aproximações vorazes e sucessivas de um sinal. Seja um dicionário $\mathcal{D} = \{\mathbf{g}_{\gamma}\}$, $\gamma \in \Gamma$, (γ é um conjunto de parâmetros que definem \mathbf{g} e Γ o conjunto de todos os possíveis γ), tal que $\|\mathbf{g}_{\gamma}\| = 1$. Para encontrar os coeficientes α_n (ver eq. (1)) e os índices $\gamma(n)$ (correspondentes a i_n na eq. (1)), escolhe-se $\mathbf{g}_{\gamma(1)}$ tal que $\alpha_1 = \max_{\gamma \in \Gamma} |\langle \mathbf{x}, \mathbf{g}_{\gamma} \rangle|$ e divide-se \mathbf{x} em duas partes, definindo o resíduo $\mathbf{r}_{\mathbf{x}}^1 = \mathbf{x} - \alpha_1 \mathbf{g}_{\gamma(1)}$. Este processo é aplicado recursivamente ao resíduo obtido, obtendo-se os coeficientes α_n e os correspondentes $\gamma(n)$ e $\mathbf{r}_{\mathbf{x}}^n$. Tem-se, então, $\mathbf{x} = \sum_{n=1}^M \alpha_n \mathbf{g}_{\gamma(n)} + \mathbf{r}_{\mathbf{x}}^M$.

A. Dicionários a partir de Frames

Para ser capaz de representar qualquer sinal um dicionário deve ser completo. Frames são bases sobrecompletas. Entretanto, *a priori*, um dicionário em $L^2(\mathbb{R})$ não é necessariamente um frame. Porém, num espaço vetorial \mathbb{R}^N , qualquer dicionário \mathcal{D} completo e finito, i.e. \mathcal{D} é tal que $\mathcal{K}(\mathcal{D}) < \infty$, usado para decompor um vetor $\mathbf{x} \in \mathbb{R}^N$ é um frame de \mathbb{R}^N . Seja um dicionário \mathcal{D} composto de elementos \mathbf{g}_k . Suponha agora que os elementos de \mathcal{D} não satisfazem a definição de frame da eq. (5). Logo, para um dado $\mathbf{x} \in \mathbb{R}^N$ teremos ou $0 = \sum_{k=1}^{\mathcal{K}(\mathcal{D})} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2$ ou $\sum_{k=1}^{\mathcal{K}(\mathcal{D})} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 = \infty$. A primeira igualdade é absurda pois, se \mathcal{D} é um dicionário em \mathbb{R}^N , para todo \mathbf{x} devemos ter ao menos um elemento $\mathbf{g}_k \in \mathcal{D}$ tal que $\langle \mathbf{x}, \mathbf{g}_k \rangle \neq 0$, caso contrário \mathbf{x} não poderia ser decomposto neste dicionário. A segunda também o é: se $\|\mathbf{x}\|^2 < \infty$ e todos os elementos \mathbf{g}_k possuem energia limitada, por exemplo, por g^2 , então necessariamente $\sum_{k=1}^{\mathcal{K}(\mathcal{D})} |\langle \mathbf{x}, \mathbf{g}_k \rangle|^2 \leq \mathcal{K}(\mathcal{D}) \|\mathbf{x}\|^2 g^2 < \infty$. Desta forma, tem-se que todo dicionário finito de \mathbb{R}^N é um frame de \mathbb{R}^N . Semelhantemente, todo frame de \mathbb{R}^N é um dicionário de \mathbb{R}^N .

Em algoritmos de decomposição vorazes, alguns dos principais fatores que determinam tanto o erro máximo quanto a velocidade de convergência da decomposição são:

- O ângulo máximo entre qualquer sinal do espaço e seu elemento mais próximo do dicionário

$$\Theta(\mathcal{D}) = \arccos \left(\min_{\mathbf{x}} \left[\max_{\gamma} \left(\frac{|\langle \mathbf{x}, \mathbf{g}_{\gamma} \rangle|}{\|\mathbf{x}\|} \right) \right] \right). \quad (18)$$

Este fator é relevante. Pois, por exemplo, usando o MP pode-se mostrar que o erro $\mathbf{r}_{\mathbf{x}}^M$, após M passos, é tal que

$\|\mathbf{r}_x^M\| \leq \|\mathbf{x}\| \sin^M(\Theta(\mathcal{D}))$. Entretanto, não há solução conhecida para o cálculo de $\Theta(\mathcal{D})$ dado um \mathcal{D} qualquer.

- O ângulo médio entre os vetores do espaço e seus elementos mais próximos do dicionários

$$\bar{\Theta} = E \left\{ \arccos \left[\max_{\gamma} \left(\frac{|\langle \mathbf{x}, \mathbf{g}_{\gamma} \rangle|}{\|\mathbf{x}\|} \right) \right] \right\}. \quad (19)$$

Ambos, $\Theta(\mathcal{D})$ e $\bar{\Theta}$, são medidas de desempenho de dicionários que mensuram, respectivamente, quanto do resíduo pode ser aproximado no passo no pior caso e na média.

Os *frame bounds* são medidas de redundância do *frame* [1]. Suponha um *frame* normalizado, isto é, $\|\mathbf{g}_k\| = g, \forall k$, para todo sinal \mathbf{x} teremos $A\|\mathbf{x}\|^2 \leq \|\mathbf{x}\|^2 g^2 \sum_k |\cos(\angle \mathbf{x}, \mathbf{g}_k)|^2 \leq B\|\mathbf{x}\|^2$, onde $\angle \mathbf{x}, \mathbf{g}_k$ é o ângulo entre \mathbf{x} e \mathbf{g}_k . Teremos então que $\frac{A}{g^2} \leq \sum_k |\cos(\angle \mathbf{x}, \mathbf{g}_k)|^2 \leq \frac{B}{g^2}$. Quanto maior $\sum_k |\cos(\angle \mathbf{x}, \mathbf{g}_k)|^2$, maior é a concentração de elementos do dicionários em direções do espaço similares à de \mathbf{x} . Quanto mais próximos forem A e B , então, teremos uma “densidade” mais parecida do *frame* em todas as direções do espaço \mathbb{R}^N . Entretanto, este argumento não pode ser tomado no sentido de optimalidade do dicionário, que é dada por $\Theta(\mathcal{D})$, já que ele não implica um $\Theta(\mathcal{D})$ ótimo, simplesmente indica uma aproximação razoável para todos os $\mathbf{x} \in \mathbb{R}^N$.

Da seção anterior, conclui-se que o método de intercalação proposto é eficiente no sentido de: (i) minimizar o produto interno entre os elementos acrescentados ao dicionário (*frame*) e os anteriormente existentes; e (ii) gerar *frames* mais apertados. (i) é relevante para a construção de dicionários para algoritmos de decomposição vorazes de sinais, pois, nestes, não basta simplesmente aumentar $\mathcal{K}(\mathcal{D})$ de qualquer forma, é interessante aumentá-la inserindo novas estruturas nos pontos mais distantes dos elementos que já estão presentes no dicionário. Desta forma, sinais que eram mal-decompostos podem ser melhor decompostos (com menos termos ou passos). Esta abordagem permite, ainda, que se obtenham dicionários que continuam parametrizáveis e com os elementos possuindo uma “forma” ou características pré-definidas.

B. Avaliação dos Dicionários Gerados

Para avaliar o desempenho obtido pelos dicionários gerados em algoritmos vorazes, iremos avaliar tanto $\Theta(\mathcal{D})$ como $\bar{\Theta}$, assim como o histograma do ângulo, isto é $\arccos \left[\max_{\gamma} \left(\frac{|\langle \mathbf{x}, \mathbf{g}_{\gamma} \rangle|}{\|\mathbf{x}\|} \right) \right]$ para diversos dicionários. Para isto, geramos diferentes dicionários a partir de *frames* WH e uma união usando o método proposto. Os dicionários gerados são apresentados nas duas primeiras colunas da Tab. I. Estes foram gerados a partir de uma Gaussiana $\mathbf{g}(k) = e^{\frac{k^2}{2\sigma^2}}$ com suporte em $[0, (N-1)]$ e extensão periódica, no espaço \mathbb{R}^{64} ($N = 64$), utilizamos, para compor o dicionário, as partes reais e imaginárias dos elementos (isto gera um *frame* com os mesmos *frame bounds*). Os resultados em termos de $\Theta(\mathcal{D})$ e $\bar{\Theta}$ apresentados na Tab. 1 são para dois valores distintos de σ^2 – uma Gaussiana estreita e outra larga. As estruturas geradoras destes dicionários (ver segunda coluna da Tab. 1) são descritas a seguir. Os dicionários A e B são respectivamente $E_{mb}T_{na}\mathbf{g}(k)$ e $E_{mb+b/2}T_{na+a/2}\mathbf{g}(k)$, ou seja as estruturas de B são “posicionadas” nos pontos de mínimos de WD_A e vice-versa. Os valores de a e b foram escolhidos de forma

TABELA I

AVALIAÇÃO DOS DICIONÁRIOS GERADOS A PARTIR DE *frames*. FORMADO A PARTIR DE UMA GAUSSIANA EM \mathbb{R}^{64} – ESTIMADOS A PARTIR DE 200 DECOMPOSIÇÕES DE 1000 PASSOS DE RÚIDO GAUSSIANO USANDO O MP.

Dicionários			$\sigma^2 = 250$		$\sigma^2 = \frac{\sqrt{2}\pi}{2}$	
\mathcal{D}	Frame Gerador	$\mathcal{K}(\mathcal{D})$	$\Theta(\mathcal{D})$	$\bar{\Theta}$	$\Theta(\mathcal{D})$	$\bar{\Theta}$
A	$E_{mb}T_{na}\mathbf{g}(k)$	256	89.3	87.6	82.4	79.3
B	$E_{mb+b/2}T_{na+a/2}\mathbf{g}(k)$	256	89.5	88.7	82.8	79.9
C	$A \cup B$	512	78.0	73.6	75.0	70.1
D	$E_{mb}T_{na/2}\mathbf{g}(k)$	512	88.3	86.4	75.5	70.0
E	$E_{mb/2}T_{na}\mathbf{g}(k)$	512	77.9	73.6	82.4	79.3
F	$E_{mb/2}T_{na/2}\mathbf{g}(k)$	1024	77.9	73.6	75.4	69.8
G	$E_{mb+b/2}T_{na/2+a/2}\mathbf{g}(k)$	512	89.4	88.7	75.4	70.1
H	$E_{mb/2+b/2}T_{na+a/2}\mathbf{g}(k)$	512	78.3	73.8	82.6	79.3

a fornecer a mesma resolução nos eixos do tempo e da frequência para os dicionários A e B (no caso, 16 pontos em cada direção, o que acarreta $a = N/16$ e $b = \pi/16$). Isto permite a mesma cardinalidade para A e B e a “dobro” de cardinalidades dos dicionários descritas a seguir. O dicionário C é formado pela união de A e B . Os dicionários D e E foram gerados de forma a possuir a mesma estrutura de A , porém dobrando sua cardinalidade de forma a serem iguais à de C . Para gerar D , o parâmetro a foi dividido por dois; e, para gerar E , b foi dividido por dois. O dicionário F foi obtido dividindo os parâmetros de deslocamento e modulação por dois, sendo então gerado por $E_{mb/2}T_{na/2}\mathbf{g}(k)$ o que implica uma “densidade” e cardinalidade 4 vezes maior em relação a A e B . Os dicionários G e H foram construídos de forma a dobrar a cardinalidade do dicionário B . Para isto a foi dividido por 2 para gerar G e b foi dividido por dois para gerar H . Todos os dicionários gerados tiveram seus elementos normalizados. A Fig. 3 apresenta os histogramas dos ângulos entre os resíduos e os elementos mais próximos dos dicionários com $\sigma^2 = 250$; na Tab. I, conforme já mencionado, temos estimativas de $\Theta(\mathcal{D})$ e $\bar{\Theta}$.

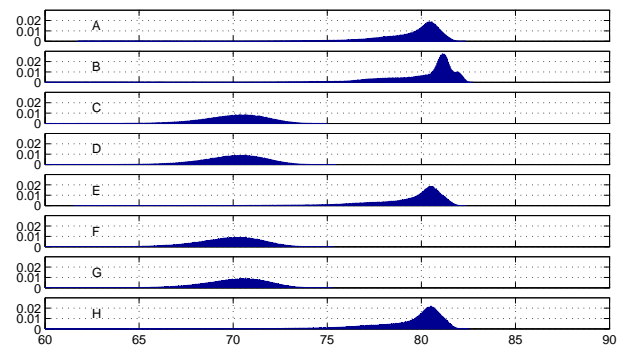


Fig. 3. Histogramas (com 500 intervalos) dos ângulos entre um sinal qualquer e o elemento mais próximo do dicionário para os dicionários A, B, C, D, E, F, G e H com $\sigma^2 = \frac{\sqrt{2}\pi}{2}$ – obtidos a partir de 200 decomposições, de sinais uniformemente distribuídos na superfície da hipersfera, com 1000 passos cada no MP – i.e. 200000 amostras do ângulo.

Já fora observado que a tendência é de que quando a cardinalidade de \mathcal{D} aumenta $\Theta(\mathcal{D})$ tende a diminuir [14], [22]. Entretanto, não é trivial saber como se dará este comportamento ao adicionar elementos a um dicionário [23]. Isto decorre de que para uma diminuição significativa de $\Theta(\mathcal{D})$ devemos posicionar os novos elementos do dicionário exatamente nas direções dos sinais que acarretam as menores taxas de aproximação, ou seja, que possuem maior ângulo

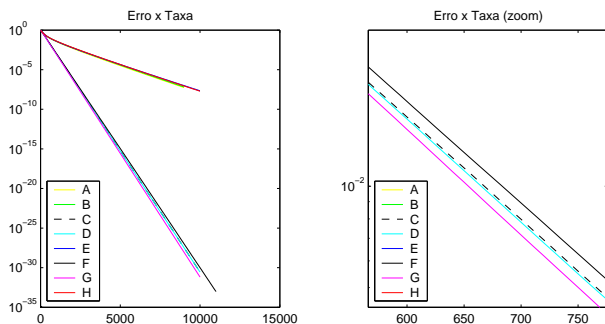


Fig. 4. Erro médio em função da taxa (passos $\times [\log_2 [\mathcal{K}(\mathcal{D})]]$) para os dicionários A, B, C, D, E, F, G e H com $\sigma^2 = \frac{\sqrt{2\pi}}{2}$ – a partir de 200 decomposições, usando MP de 1000 passos, de sinais uniformemente distribuídos na superfície da hipersfera.

com o elemento do dicionário mais próximo. Os resultados apresentados, tanto na Tab. I quanto na Fig. 3, mostram que ao posicionar os novos elementos de um dicionário nos mínimos do conteúdo tempo-frequência do dicionário (ou do *frame* correspondente) gerado a partir de um *frame* de Gabor é mais eficiente para diminuir $\Theta(\mathcal{D})$ e $\bar{\Theta}$ do que simplesmente dividir a ou b por dois. Na verdade, dos resultados podemos observar que o posicionamento proposto para os novos elementos, obtém um dicionário com $\Theta(\mathcal{D})$ e $\bar{\Theta}$ mais próximos aos obtidos ao dividir a e b por dois (abordagem do dicionário F , com a metade da cardinalidade (o que implica menor complexidade computacional), do que as outras abordagens. A Fig. 4 apresenta o erro médio (em escala logarítmica) em função da “taxa” (obtida multiplicando o número de passos por $[\log_2 [\mathcal{K}(\mathcal{D})]]$) para os oito dicionários com $\sigma^2 = \frac{\sqrt{2\pi}}{2}$. Apesar de similares os melhores desempenhos em termos de taxa-distorção são obtidos, neste caso, pelos dicionários C, D, G e F . Experiências realizadas alterando: a largura da Gaussiana (σ^2), os parâmetros a, b e a fase dos elementos dos dicionários, às vezes geram dicionários com as estruturas de D, E, G ou H com desempenhos próximos ao de C e F . Como exemplo, veja as duas colunas para $\sigma^2 = 250$ da Tab. I, nas quais vemos que os dicionários com as estruturas de E e H possuem medidas de desempenho ($\Theta(\mathcal{D})$ e $\bar{\Theta}$) similares às de C , e para $\sigma^2 = \frac{\sqrt{2\pi}}{2}$ os dicionários com as estruturas de D e G possuem medidas de desempenho similares às de C . Entretanto, nessas experiências vimos que a estrutura proposta para C sempre gera dicionários com medidas de desempenho similares às obtidas com a abordagem usada para gerar F (que possui cardinalidade 2 vezes maior). Isto posto, o método de aumento controlado da cardinalidade do dicionário (e portanto do custo computacional da decomposição) por intercalação sempre produz bons resultados. Este é o ponto relevante: o compromisso entre $\mathcal{K}(\mathcal{D})$ e $\bar{\Theta}$ que o método apresentado atinge de forma satisfatória para a adição de elementos a dicionários parametrizáveis.

VI. CONCLUSÕES

Neste trabalho, baseado no conteúdo tempo-frequência dos elementos de um *frame*, uma condição suficiente que garante que um conjunto de elementos g_k qualquer gera um *frame* foi apresentada. Apresentou-se também uma discussão detalhada do conteúdo tempo-frequência de *frames* de Weyl-Heisenberg gerados a partir de funções pares e ímpares. Os resultados

sugeriram um método de obter *frames* cada vez mais apertados. Este resultado permite uma abordagem para aumentar a cardinalidade de dicionários utilizando a intercalação de *frames* no espaço de parâmetros. Os dicionários assim gerados foram avaliados no algoritmo de decomposição voraz conhecido como *matching pursuits*. O método de união de dicionários por intercalação no espaço de parâmetros mostrou-se eficiente para obter dicionários parametrizáveis em função de seus deslocamentos no tempo e na frequência, construídos a partir de uma função $g(t)$. Os dicionários assim obtidos são capazes de obter decomposições com média da norma do erro menor, para uma dada cardinalidade.

REFERÊNCIAS

- [1] S. Mallat, *A Wavelet Tour of Signal Processing*, 1st ed. San Diego, California, USA: Academic Press, 1998.
- [2] H. Krim, D. Tucker, S. Mallat, and D. Donoho, “On denoising and best signal representations,” *IEEE Trans. on Information Theory*, vol. 45, pp. 2225–2238, 1999.
- [3] K. Egan, S. O. Aase, and J. H. Husoy, “Multi-frame compression: Theory and design,” *Elsevier Signal Processing*, vol. 80, pp. 2121–2140, 2000.
- [4] L. Lovisolo, M. A. M. Rodrigues, E. A. B. da Silva, and P. S. R. Diniz, “Efficient coherent decompositions of power systems signals using damped sinusoids,” *IEEE Trans. on Signal Processing*, 2005, Aceito para publicação.
- [5] C. E. Etemoglu and V. Cuperman, “Matching pursuit sinusoidal speech coding,” *IEEE Trans. on Speech and Audio Processing*, vol. 11, pp. 413–424, 2003.
- [6] P. Vera-Candes, N. Ruiz-Reyes, M. Rosa-Zurera, and D. Martinez-Munoz, “Transient modeling by matching pursuits with a wavelet dictionary for parametric audio coding,” *IEEE Signal Processing Letters*, vol. 11, pp. 349–352, 2004.
- [7] R. Caetano, E. A. B. da Silva, and A. G. Ciancio, “Matching pursuits video coding using generalized bit-planes,” in *IEEE Inter. Conf. on Image Processing*, Setembro 2002.
- [8] R. Gribonval and E. Bacry, “Harmonic decomposition of audio signals with matching pursuit,” *IEEE Trans. on Signal Processing*, vol. 51, pp. 101–111, 2003.
- [9] R. Heusdens, R. Vafin, and W. B. Kleijn, “Sinusoidal modeling using psychoacoustic-adaptive matching pursuits,” *IEEE Signal Processing Letters*, vol. 9, pp. 262–265, 2002.
- [10] M. M. Goodwin and M. Vetterli, “Matching pursuits and atomic signal models based on recursive filters banks,” *IEEE Trans. on Signal Processing*, vol. 47, pp. 1890–1902, 1999.
- [11] S. Jaggi, W. Carl, S. Mallat, and A. S. Willsky, “High resolution pursuit for feature extraction,” *Applied and Computational Harmonic Analysis*, vol. 5, pp. 428–449, 1998.
- [12] S. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Trans. on Signal Processing*, vol. 41, pp. 3397–3415, 1993.
- [13] J. Adler, B. D. Rao, and K. Kreutz-Delgado, “Comparison of basis selection methods,” in *Proceedings of the 30th Asilomar Conf. on Signals, Systems and Computers*, Novembro 1996, pp. 252–257.
- [14] G. Davis, “Adaptive nonlinear approximations,” Ph. D. Thesis, Courant Institute of Mathematical Sciences, Computer Science Department, New York University, 1994.
- [15] K. Egan, S. O. Aase, and J. H. Husoy, “Designing frames for matching pursuits algorithms,” in *IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing*, Maio 1998, pp. 1817–1820.
- [16] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, Pennsylvania, USA: Society for Industrial and Applied Mathematics, 1991.
- [17] O. Christensen, *An Introduction To Frames And Riesz Bases*, 1st ed., ser. Applied and Numerical Harmonic Analysis. Boston, Basel, Berlin: Birkhäuser, 2002.
- [18] Y. C. Eldar and G. D. Forney Jr., “Optimal tight frames and quantum measurement,” *IEEE Trans. on Information Theory*, vol. 48, pp. 599–610, 2002.
- [19] Y. C. Eldar and H. Bölcskei, “Geometrically uniform frames,” *IEEE Trans. on Information Theory*, vol. 49, pp. 993–1006, 2003.
- [20] L. Lovisolo, M. G. Pinho, E. A. B. da Silva, and P. S. R. Diniz, “On the time-frequency content of Weyl-Heisenberg frames generated from odd and even functions,” in *IEEE Inter. Conf. on Circuits and Systems*, Maio 2005.
- [21] T. Strohmer and R. W. Heath Jr., “Grassmannian frames with applications to coding and communications,” *Applied and Computational Harmonic Analysis*, vol. 14, pp. 257–275, 2003.
- [22] Q. Liu, Q. Wang, and L. Wu, “Size of the dictionary in matching pursuit algorithm,” *IEEE Trans. on Signal Processing*, vol. 52, pp. 3403–3408, 2004.
- [23] R. Caetano, “Codificação de vídeo usando planos de bits generalizados,” D. Sc. Thesis, COPPE/UFRJ, Rio de Janeiro, RJ, 2004.